



TITLE:

Improving Mutual Understanding in Machine Translation Mediated Communication(Abstract_要旨)

AUTHOR(S):

Mondheera, Pituxcoosuvarn

CITATION:

Mondheera, Pituxcoosuvarn. Improving Mutual Understanding in Machine Translation Mediated Communication. 京都大学, 2020, 博士(情報学)

ISSUE DATE:

2020-03-23

URL:

<https://doi.org/10.14989/doctor.k22579>

RIGHT:

博士學位論文調查報告書

論文題目 Improving Mutual Understanding in Machine Translation Mediated
Communication
(機械翻訳を介したコミュニケーションにおける相互理解の改善)

申請者氏名 MONDHEERA PITUXCOOSUVARN

最 終 学 歴 平成 28年 3月
京都大学大学院情報学研究科社会情報学専攻修士課程 修了
令和 2年 3月
京都大学大学院情報学研究科社会情報学専攻博士後期課程
研究指導認定見込

学 識 確 認 令和 年 月 日（論文博士のみ）

論文調査委員 京都大学大学院情報学研究科
(調査委員長) 教 授 緒方 広明

論文調査委員 京都大学大学院情報学研究科
特定准教授 林 冬恵

論文調査委員 京都大学大学院情報学研究科
教授 河原 達也

論文調査委員 京都大学
名誉教授 石田 亨

(続紙 1)

京都大学	博士（情報学）	氏名	MONDHEERA PITUX COOSUVARN
論文題目	Improving Mutual Understanding in Machine Translation Mediated Communication （機械翻訳を介したコミュニケーションにおける相互理解の改善）		
（論文内容の要旨）			
<p>The goal of this study is to improve mutual understanding in machine translation (MT) mediated communication where the users have different languages and cultural backgrounds. There are several typical issues that hinder mutual understanding in traditional MT mediated communication where users use their mother languages and MT is used to translate users' messages. First, when the qualities of MT for each language are different, the users who speak different languages have an unequal opportunity to understand the conversation. Second, users with different cultural backgrounds sometimes perceive culturally related concepts differently and fail to establish a common ground for communication consequently. To cope with the above issues, this thesis presents methods to create novel communication environments that improve users' mutual understanding. This thesis consists of seven chapters.</p> <p>Chapter 1 outlines the thesis including the objectives of this research, issues, and approaches.</p> <p>Chapter 2 describes the problems and solutions in MT mediated communication. In order to understand the real-world communication problems when MT is used, this chapter introduces an ethnographic study on an intercultural children workshop where the children from different countries worked together and communicated via MT. We found that the users who speak low-resource languages faced difficulties to understand mistranslated messages and switched to different meanings of communication including using English for a better understanding of the conversation. In addition, misunderstandings occurred when users from different countries had different understandings regarding culturally related concepts; however, they solved this problem by searching for images on the Internet to ground their understandings of these concepts in different cultures. The observed problems and their solutions in the real field motivated our research in this thesis.</p> <p>Chapter 3 proposes a model of best-balanced machine translation (BBMT) that allows the users to have an equal opportunity to participate and understand the conversation. The purpose of BBMT is to select a best-balanced communication channel in MT mediated communication, where each channel indicates a set of users' languages and MT services. In the BBMT model, the quality of message for each possible communication channel is first calculated by using users' language scores and MT quality. Then, the best-balanced communication channel can be selected by searching for the channel with Pareto optimal quality of message. An experiment on communication in collaborative games shows that the users are able to participate more equally and face fewer conversational breakdowns by using BBMT compared to the cases using traditional MT mediated communication and using English as a shared language. This result indicates that BBMT can lower the misunderstanding and improve mutual understanding during the communication.</p> <p>Chapter 4 presents a model of privacy-aware calculation for BBMT by addressing the issue that the original BBMT requires users' language scores to be disclosed. When using BBMT in MT mediated communication, it is important to treat language scores as users' private data and protect the privacy from being disclosed. Existing methods for privacy protection mainly focus on the data itself and therefore are difficult to be applied for BBMT which requires several</p>			

steps of calculation and selection. The objective here is to enable BBMT without disclosing users' language scores. To this end, secure computation technologies and multi-agent systems are applied to establish a model for protecting users' privacy. The proposed model divides the tasks of calculation and selection in BBMT into distributed pieces and assigns them to different agents so that none of the agents retains enough information to know or guess the private data, in this case, users' language scores. Security arguments are used to prove that the proposed model can successfully protect users' language scores while making the tasks of calculation and selection possible in BBMT.

Chapter 5 provides an automated cultural difference detection method for supporting users' mutual understanding in MT mediated communication. In intercultural communication, misunderstandings on culturally related concepts frequently occur when users have different cultural backgrounds. On the other hand, images can help users ground their understandings of culturally related concepts in different cultures. Since there are links among concepts, words, and images on the Internet, an automated method is proposed to detect cultural differences by using concept dictionaries and image similarity calculation. First, words in different languages are extracted from concept dictionaries, and are used to search for the corresponding images. Then, the similarity of the retrieved images is automatically calculated to detect the cultural difference of the words. An experiment in which Japanese WordNet and English WordNet are used as concept dictionaries shows that the proposed method can successfully find the concepts that cause misunderstandings in MT mediated communication.

Chapter 6 proposes a system that uses BBMT and automated cultural difference detection to improve mutual understanding in MT mediated communication. To confirm the effect of the system, this chapter also reports a controlled experiment on collaborative problem solving, where the controlled group members use traditional MT mediated communication while the experimental group members use BBMT with the warning from the automated cultural difference detection. The result shows that the experimental group members can fully understand the messages and have less incomprehensible MT-translated messages reported, which indicates that the users can improve mutual understanding by using BBMT and automated cultural difference detection.

Chapter 7 concludes the thesis by summarizing the contributions to improve mutual understanding in MT mediated communication and suggesting possible future directions.

(論文審査の結果の要旨)

本研究は、機械翻訳を介したコミュニケーションにおける参加者の相互理解を改善することを目的としている。そのために、参加者の発話機会の均等化、参加者のプライバシーの保護、参加者の文化依存概念に対する理解の支援という代表的な課題を解決する手法を提案している。得られた主要な成果は以下の通りである。

1. 参加者の発話機会を均等化するための最適均衡機械翻訳

従来の機械翻訳を介したコミュニケーションでは、参加者が母語を利用するため、機械翻訳の品質が低下することによって、会話の理解が困難になる。一方、参加者の使用言語を英語に統一すると、非母語話者は著しく不利になる。そこで、参加者が言語的に不利にならないように、参加者の発話機会の均等化のための最適均衡機械翻訳を提案している。具体的には、参加者の言語能力と機械翻訳の品質を用いて通信路の品質を定式化し、最適な通信路を選択することによって、参加者の使用言語と参加者を繋ぐ機械翻訳の利用を最適に制御する。実験によって、既存手法に比べて参加者の発言機会の均等化を改善できることを示している。

2. 参加者のプライバシーを保護するためのマルチエージェント秘匿計算

最適均衡機械翻訳を用いたコミュニケーションでは、参加者の言語能力などのプライバシーを開示しないようシステムを設計する必要がある。従来のプライバシー保護に関する手法は、複雑なタスクを含む最適均衡機械翻訳に適用することが困難である。そこで、マルチエージェントに基づく秘匿計算手法を提案している。具体的には、最適均衡機械翻訳におけるタスクを複数の参加者エージェントと計算エージェントに分散し、秘匿計算を用いることで単体のエージェントが参加者のプライバシーを取得できないようなプロトコルを実現する。また、提案する秘匿計算の性質に関する分析によって、参加者のプライバシーを開示せずに最適均衡機械翻訳を実現可能にすることを確認している。

3. 参加者の文化依存概念に対する理解を支援するための文化差の自動検出

異文化コミュニケーションでは、仮に概念が適切に表現されていたとしても、文化差によりその意味を理解できないという問題が生じる場合がある。そこで、文化差のある概念を自動的に検出する手法を提案している。具体的には、概念辞書を用いて多言語の単語を抽出し、それらの単語を用いて画像を検索し、画像の類似度計算によって文化差のある概念を自動的に検出する。実験によって、機械翻訳を介したコミュニケーションにおいて、文化差のある概念を自動的に検出し提示することで、参加者の相互理解を改善できることを明らかにしている。

以上、本論文は、参加者の発話機会の均等化、参加者のプライバシーの保護、参加者の文化依存概念に対する理解の支援を実現する上で有効な手法を提案し、言語と文化を異にする参加者の相互理解を改善するコミュニケーション環境の構築に寄与している。よって、本論文は博士(情報学)の学位論文として価値あるものと認める。また、令和2年2月18日に実施した論文内容とそれに関連した試問の結果合格と認めた。また、本論文のインターネットでの全文公表についても支障がないことを確認した。

要旨公開可能日： 年 月 日以降